Random variables with moment-matching staircase density functions

Luis G. Crespo [a,*], Sean P. Kenny [a], Daniel P. Giesy [a], Bret K. Stanford [b]

[a] Dynamic Systems and Controls Branch, NASA Langley Research Center, Hampton, VA 23681, USA
[b] Aeroelasticity Branch, NASA Langley Research Center, Hampton, VA 23681, USA

A B S T R A C T

This paper proposes a family of random variables for uncertainty modeling. The variables of interest have a bounded support set, and prescribed values for the first four moments. We present the feasibility conditions for the existence of any of such variables, and propose a class of variables that conforms to such constraints. This class is called staircase because the density of its members is a piecewise constant function. Convex optimization is used to calculate their distributions according to several optimality criteria, including maximal entropy and maximal log-likelihood. The flexibility and efficiency of staircases enable modeling phenomena having a possibly skewed and/or multimodal response at a low computational cost. Furthermore, we provide a means to account for the uncertainty in the distribution caused by estimating staircases from data. These ideas are illustrated by generating empirical staircase predictor models. We consider the case in which the predictor matches the sample moments exactly (a setting applicable to large datasets), as well as the case in which the predictor accounts for the sampling error in such moments (a setting applicable to sparse datasets). A predictor model for the dynamics of an aeroelastic airfoil subject to flutter instability is used as an example. The resulting predictor not only describes the system's response accurately, but also enables carrying out a risk analysis for safe flight.

1. Introduction

Metamodeling [1] is the process of creating a mathematical model of a phenomenon based on given data. Making a prediction match the observations by adjusting the hyper-parameters of a computational model is a long standing approach in Bayesian inference, reliability-based design optimization, moment matching algorithms, and backward propagation of variance [2–5].

This paper proposes a framework for characterizing a univariate random variable subject to moment constraints. These constraints might stem from matching sample moments. A multitude of moment problems and applications are available in the literature. For example, moment inequalities are used to provide robust estimates for financial quantities [6,7], wealth balance, and portfolio optimization [8]. In the decision science area [9], moment-bounds are used for dynamic programming, decision analysis with incomplete information, and Bayesian statistics.

* Corresponding author.
E-mail address: luis.g.crespo@nasa.gov (L.G. Crespo).

In particular, we focus on variables having a finite support set and prescribed values for the first four moments. The interdependency among the requirements associated with this prescription determines the feasibility conditions for the existence of any of such variables. We first study this feasible space, and then chose a particular class of variables from the possibly infinitely many that satisfy the constraints. This class is called *staircase* because the density of its members is a piecewise constant function on a finite collection of subintervals partitioning the chosen support set. Staircases enable describing a wide range of density shapes, including multimodal and/or strongly skewed distributions. In contrast to the standard spectrum of distributions, however, a staircase variable does not admit an algebraic representation but instead it is determined by solving an optimization program.

Staircase variables can be shaped according to one of several optimality criteria including maximal entropy, minimal squared amplitude, optimal target matching, and maximal log-likelihood. Convexity enables substantial computational advantages including fast computational times, and the ability to solve problems with large number of design variables and constraints using interior point methods. This feature makes them well suited for many practical modeling applications requiring their repeated calculation and simulation, e.g., the Bayesian calibration of a computational model whose parameters are staircases. Furthermore, we develop a technique that enables accounting for the sampling error that results from basing the estimate of a staircase distribution on a finite set of observations. The modeling and risk analysis of the dynamics of an aeroelastic airfoil subject to flutter instability is used as an example.

2. Preliminaries

Consider the continuous random variable z with support set $\Delta_z = [z_{\min},\ z_{\max}]$, and density function $f_z : \Delta_z \subset \mathbb{R} \to \mathbb{R}^+$. Denote by m_r the rth central moment of z, which is defined as

$$m_r = \int_{\Delta_z} (z-\mu)^r f_z(z)dz,\ r = 0, 1, 2, \ldots \tag{1}$$

where μ is the expected value of z. Note that $m_0 = 1,\ m_1 = 0,\ m_2$ is the variance, m_3 is the third-order central moment, and m_4 is the fourth-order central moment. In this paper, when reference is made to the rth moment of a random variable, we assume that the corresponding integral in (1) converges for that distribution.

The random variables of interest are constrained to have a bounded support set and given values for μ, m_2, m_3, and m_4. The bounded support constraint is given by $\Delta_z \subseteq \Omega_z$, where $\Omega_z = [\underline{z}, \bar{z}]$, with $\bar{z} \geq \underline{z}$. The moment constraints are the equality constraints in (1). The parameters of these constraints will be grouped into the variable $\theta_z \in \mathbb{R}^6$ given by

$$\theta_z = [\underline{z}, \bar{z}, \mu, m_2, m_3, m_4]. \tag{2}$$

3. Feasibility of a variable constrained by θ

A set of sufficient conditions for the feasibility of a random variable z satisfying the constraints associated with θ_z, derived from the work of Sharma [10,11] and Kumar [12], is given by[1] $g(\theta_z) \leq 0$, where

$$\begin{aligned}
g_1 &= \underline{z} - \bar{z},\\
g_2 &= \underline{z} - \mu,\\
g_3 &= \mu - \bar{z},\\
g_4 &= -m_2,\\
g_5 &= m_2 - v,\\
g_6 &= m_2^2 - m_2(\mu - \underline{z})^2 - m_3(\mu - \underline{z}),\\
g_7 &= m_3(\bar{z} - \mu) - m_2(\bar{z} - \mu)^2 + m_2^2,\\
g_8 &= 4m_2^3 + m_3^2 - m_2^2(\bar{z} - \underline{z})^2,\\
g_9 &= 6\sqrt{3}m_3 - (\bar{z} - \underline{z})^3,\\
g_{10} &= -6\sqrt{3}m_3 - (\bar{z} - \underline{z})^3,\\
g_{11} &= -m_4,\\
g_{12} &= 12m_4 - (\bar{z} - \underline{z})^4,\\
g_{13} &= (m_4 - vm_2 - um_3)(v - m_2) + (m_3 - um_2)^2,\\
g_{14} &= m_3^2 + m_2^3 - m_4m_2,
\end{aligned}$$

for $u = \bar{z} + \underline{z} - 2\mu$ and $v = (\mu - \underline{z})(\bar{z} - \mu)$. Feasibility conditions for variables supported in either $\Delta_z = \mathbb{R}$ or $\Delta_z = \mathbb{R}^+$ having moments of arbitrary order can be described as positive semidefinite constraints [13,14]. Sufficient and necessary conditions

[1] Vector inequalities hold component-wise.

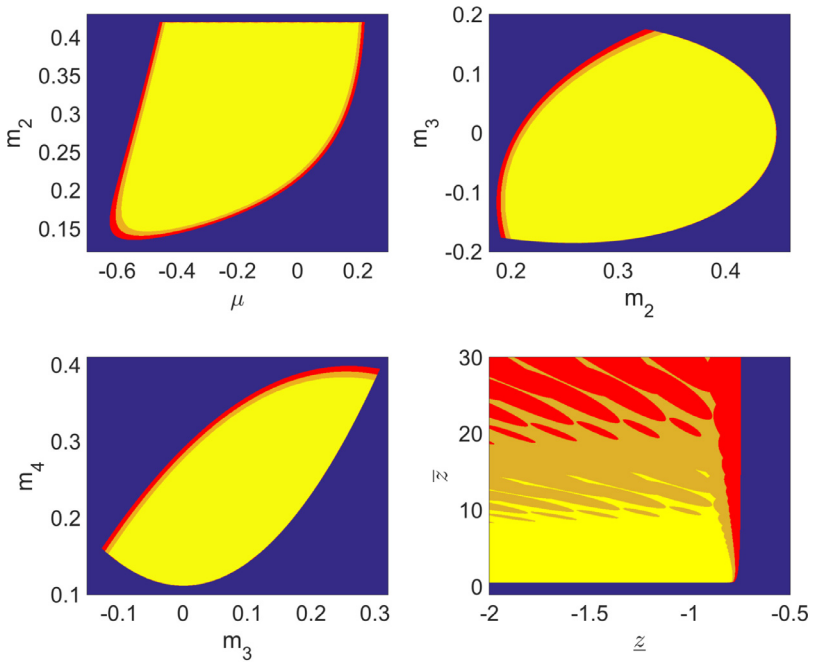

Fig. 1. Intersections of Θ with four 2-dimensional hyperplanes (non-blue region) along with $\mathcal{S}(50)$ (yellow region) and $\mathcal{S}(100)$ (union of the regions in yellow and orange). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

for moment feasibility given a finite support, called the Hausdorff moment problem, are given as a set of infinitely many constraints.

The θ-feasible domain, Θ, is defined as

$$\Theta = \{\theta : g(\theta) \leq 0\}. \tag{3}$$

It can be shown that the set Θ is non-convex. Fig. 1 shows intersections of Θ with some 2-dimensional hyperplanes. These intersections are the complement of the regions in blue. The significance of the colors will be explained in the next section. The top-left, top-right, and bottom-left subplots correspond to $\theta_z = [-1, 1, \mu, m_2, 0.1, 0.2]$, $\theta_z = [-1, 1, 0.1, m_2, m_3, 0.2]$, and $\theta_z = [-1, 1, -0.2, 1/3, m_3, m_4]$, respectively. In these three cases the resulting sets are bounded and have vertices on their boundaries. By contrast, the bottom right plot, which corresponds to $\theta_z = [\underline{z}, \bar{z}, 0.1, 1/3, -0.15, 0.2]$, yields an unbounded set with a smooth boundary.

4. Staircase random variables

There might be infinitely many random variables that realize a point in Θ. Note that most of the standard families of random variables can only realize lower-dimensional subsets of Θ; e.g., a four hyper-parameter Beta random variable can only represent points falling into the intersection of Θ with a 4-dimensional manifold. Families with six or more hyper-parameters often introduce spurious interdependencies among the components of θ they are able to realize. These dependencies might considerably restrict their ability to realize most of Θ. Staircase variables, to be introduced next, eliminate this deficiency.

Consider a random variable with density $f_{z(h)}(z)$, where h is a hyper-parameter to be prescribed. The hyper-parameter corresponding to any random variable matching the constraints implied by θ_z is given by

$$\hat{h} = \underset{h}{\operatorname{argmin}} \left\{ J(h) : \int_{\underline{z}}^{\bar{z}} z f_{z(h)}(z) dz = \mu, \int_{\underline{z}}^{\bar{z}} (z-\mu)^r f_{z(h)}(z) dz = m_r, \quad r = 0, 2, 3, 4 \right\}, \tag{4}$$

where J is an arbitrary cost function. For simplicity in the presentation we will defer the selection of J to a later paragraph in this section. The optimization in (4) is infeasible for most standard families of random variables over most of Θ.

The concept of a staircase function (or step function) is presented first. Given a partition of the real numbers into a finite number of intervals, a staircase function on that partition is a function which is constant on each interval. The use to which we will put these functions is insensitive to the value the function takes at an endpoint shared by two adjacent intervals, so we will not labor that point. We focus our attention on the family of random variables having a staircase density, which will be called *staircase* variables hereafter. Further, if we are seeking a staircase variable which realizes the parameters in θ_z, we first pick a positive integer n_b to be the number of intervals in the partition. We then require that the

interval $[\underline{z}, \bar{z}]$ be partitioned into n_b subintervals of equal length $\kappa = (\bar{z} - \underline{z})/n_b$, by the partitioning points $z_i = \underline{z} + (i-1)\kappa$ for $1 \le i \le n_b + 1$. The staircase heights are then given by $\ell \in \mathbb{R}^{n_b}$ with $\ell \ge 0$ and $\kappa \sum_{i=1}^{n_b} \ell_i = 1$. Then

$$f_{z(h)}(z) = \begin{cases} \ell_i & \forall z \in (z_i, z_{i+1}] \text{ for } 1 \le i \le n_b, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

Next we apply Eq. (4) to a staircase variable. In this setting, the hyper-parameter becomes $h = [\theta_z, n_b]$ where $\theta_z \in \Theta$. The resulting staircase variable, to be denoted as

$$z \sim S_z(\theta_z, n_b, J), \tag{6}$$

has the density function prescribed in Eq. (5), where ℓ is given by

$$\hat{\ell} = \underset{\ell \ge 0}{\operatorname{argmin}} \left\{ J(\ell) : \sum_{i=1}^{n_b} \int_{z_i}^{z_{i+1}} z \ell_i dz = \mu, \right. \tag{7}$$

$$\left. \sum_{i=1}^{n_b} \int_{z_i}^{z_{i+1}} (z-\mu)^r \ell_i dz = m_r, \quad r = 0, 2, 3, 4 \right\}.$$

Eq. (7) can be written as

$$\hat{\ell} = \underset{\ell \ge 0}{\operatorname{argmin}} \{ J(\ell) : A(\theta, n_b)\ell = b(\theta), \quad \theta \in \Theta \}. \tag{8}$$

where $A \in \mathbb{R}^{5 \times n_b}$ and $b \in \mathbb{R}^5$ are given by

$$A = \begin{bmatrix} \kappa e \\ \kappa c \\ \kappa c^2 + \kappa^3/12 \\ \kappa c^3 + \kappa^3 c/4 \\ \kappa c^4 + \kappa^3 c^2/2 + \kappa^5/80 \end{bmatrix}, \tag{9}$$

$$b = \begin{bmatrix} 1 \\ \mu \\ \mu^2 + m_2 \\ m_3 + 3\mu m_2 + \mu^3 \\ m_4 + 4m_3\mu + 6m_2\mu^2 + \mu^4 \end{bmatrix}, \tag{10}$$

where $c \in \mathbb{R}^{n_b}$ are the centers of the bins, $c_i = (z_i + z_{i+1})/2$ with $1 \le i \le n_b$, c^n is the component-wise nth power of c, and $e \in \mathbb{R}^{n_b}$ is a vector of ones.

Several cost functions J in $z \sim S_z(\theta_z, n_b, J)$ are studied next. The first cost we consider is

$$J(\ell) = -E(\ell) := \kappa \log(\ell)^\top \ell, \tag{11}$$

where E is the differential entropy. The principle of maximum entropy states that the probability distribution which best represents the current state of knowledge given some prior information is the one with largest entropy. In other words, one should select the distribution which leaves you the largest remaining information uncertainty/entropy that is consistent with the constraints. As such, no additional assumptions or biases are introduced into the calculations. In the context of staircase variables the prior information is the set of constraints associated with θ_z. Hence, the solution to (8) for the cost function in (11) yields a staircase random variable of maximal entropy.

Another class of cost functions is

$$J(\ell) = H(\ell, Q, f) := \ell^\top Q \ell + f^\top \ell, \tag{12}$$

where $Q \in \mathbb{R}^{n_b \times n_b}$ is a positive semi-definite matrix and $f \in \mathbb{R}^{n_b}$. This structure makes (7) a quadratic program. For example, using $J = H(\ell, I, 0)$ in (7), where I is the identity matrix, yields a staircase variable that minimizes the squared sum of the likelihoods at the bins. As another example, if $t \in \mathbb{R}^{n_b}$ is the density of a target random variable at c, $J = H(\ell,\ I, -2t)$ yields a staircase variable that minimizes the squared difference between the densities of the staircase and the target. This formulation provides a means to make a staircase variable conform to any particular shape while satisfying the required constraints.

Another cost function is the (logarithm of the) likelihood of the data in the sequence $\mathbb{D} = \{z^{(j)}\}$, for $j = 1, \ldots N$. When the observations in the sequence are independent we have

$$J(\ell) = -L(\ell, \mathbb{D}) := -\omega^\top \log(\ell), \tag{13}$$

where $\omega \in \mathbb{R}^{n_b}$ is given by

$$\omega_i = \sum_{j=1}^{N} \mathcal{I}\left\{ z^{(j)} \in [z_i, z_{i+1}] \right\},$$

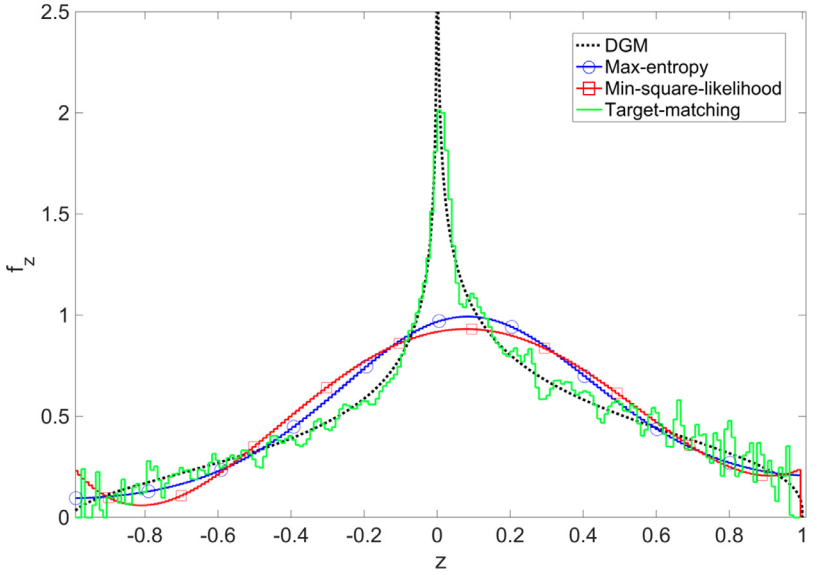

Fig. 2. Staircase densities for maximal entropy (blue), minimal squared likelihood (red) and target-matching (green) for $n_b = 200$. The density of the DGM is shown in black. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and $\mathcal{I}$ is the indicator function. Staircase variables based on (13) maximize the likelihood of the data. Maximal likelihood staircases often have PDFs with large jumps between neighboring bins. Means to smooth out such jumps are provided in the Appendix.

Staircase variables having different cost functions for a fixed hyper-parameter h are shown in Fig. 2. The components of θ are empirical estimates obtained from an ensemble of $N = 1000$ data points drawn from the *Data Generating Mechanism* (DGM) $z = u^2 - v^3$, where u and v are standard uniforms. This figure shows a maximal entropy staircase, a minimum squared likelihood staircase, and a target-matching staircase based on the empirical density. The maximal entropy variable exhibits the smoothest varying density whereas the target-matching conforms best to the DGM. The noise in the target-matching density, caused by noise in the empirical density, can be eliminated by using kernel smoothing techniques. The staircase that maximizes the likelihood of the data is studied in the Appendix.

Staircases attaining other optimality criteria, including the greatest/smallest probability of a given event[2], the maximal/minimal value of a given quantile, and the prediction intervals of greatest/smallest width can be easily formulated. Each of these pairs of problems yields a range of values where the metric of interest for all other staircase variables bounded by θ must fall. Smaller ranges are obtained when additional constraints such as symmetry, unimodality, convexity, or smoothness are imposed. By calculating the range of all possible values we eliminate the subjectivity that results from assuming a particular distribution.

When J is a convex function, the optimization program in Eq. (8) is convex so it has a unique minimum. Convexity enables substantial computational advantages including fast computational times, the ability to efficiently solve problems with a large number of design variables and constraints (on the order of hundreds of thousands), and the guarantee that numerical search converges to the global minimum. By encoding the feasible set of (8) using a computationally tractable self-concordant barrier, interior point methods guarantee that the number of iterations required to converge to the optimum is bounded by a polynomial in the dimension and accuracy of the solution [17]. Fig. 3 shows the dependency of the CPU time needed to calculate a staircase variable versus the number of bins. These calculations were performed on a standard desktop computer using commercially available software. The prescription of a random variable entails setting a hierarchy of infinitely many moments. In the case of staircases the first four moments are prescribed by constraints whereas all other moments are prescribed by the cost function. Hence, changes in J might render variables with significantly different higher-order moments, thus, shapes.

A variety of approaches to density estimation are available, including kernel density estimation [18] and a range of data clustering techniques [19]. In contrast to these approaches however, staircases do not seek to represent the density as a combination of kernel functions. This eliminates the complications the result from assuming an unsuitable basis, practice that often leads to ill-conditioned optimization programs, densities with spurious patterns, and probability tails extending beyond the intended range, e.g., representing a uniform variable as a mixture of Gaussians. Instead, the proposed

[2] Semidefinite optimization can be used to calculate tight bounds of $P[a \le z \le b]$ when $\Delta_z = \mathbb{R}$ or $\Delta_z = \mathbb{R}^+$ for arbitrary distributions having moments of arbitrary order [14]. The developments in [15,16] enable the consideration of additional structural properties such as symmetry, unimodality, convexity, and smoothness.

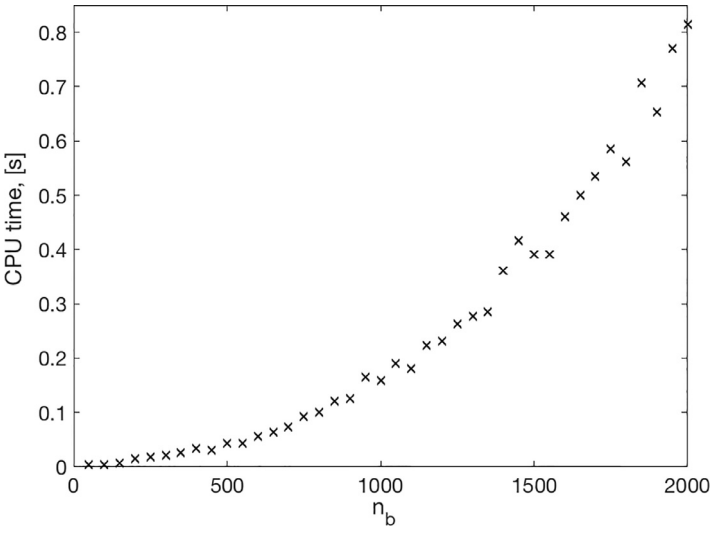

Fig. 3. CPU time versus number of bins.

formulation seeks extreme members of a moment-matching probability box using convex optimization. Staircase variables of maximal entropy are well suited for datasets of small size, whereas staircase variables of maximal (log) likelihood are better suited for datasets of moderate and large sizes. This is the result of the empirical moments converging faster to their true values than the staircase density converging to the true density. The densities of maximal likelihood variables corresponding to different datasets all having the same small number of observations will vary significantly. In this case the randomness of the sample, rather than the underlying DGM distribution, shapes the resulting variable. This observation applies to all maximal likelihood variables regardless of their structure, i.e., staircase or not.

5. Staircase feasibility

The assumed value for n_b might make (8) infeasible. This section formalizes this notion. The staircase-feasible domain is defined as[3]

$$\mathcal{S}(n_b) = \{\theta \in \Theta : \exists\, \ell \geq 0 \mid A(\theta, n_b)\ell = b(\theta)\}.$$

Hence, the set $\mathcal{S}(n_b)$ is comprised of all realizations of θ for which an staircase variable having n_b bins exists. A few comments regarding staircase feasibility are in order. Note that θ-feasibility does not imply staircase feasibility, i.e., $\theta \notin \mathcal{S}(n_b)$ even though $\theta \in \Theta$. Further notice that the chosen cost J does not affect feasibility.

Determining the membership of θ into $\mathcal{S}(n_b)$ require evaluating the feasibility of $A(\theta, n_b)\ell = b(\theta)$ for $\ell \geq 0$. Rigorous means to carry out this analysis are impractical because of their exponential dependency on n_b. In this study however, staircase feasibility is determined by solving for $\hat{\ell}$ in Eq. (8) and determining if the converged solution is a feasible minimum. This approach is computationally viable thanks to the convexity of the optimization program.

We will now revisit Fig. 1 to numerically compare $\mathcal{S}$ with Θ. The colors therein show intersections of $\mathcal{S}(50)$ (region in yellow) and $\mathcal{S}(100)$ (union of the regions in yellow and orange) with two dimensional planes in Euclidean four-space. These figures were generated by calculating a staircase variable matching the corresponding value of θ at each point in a fine grid, and then evaluating the convergence of the optimization algorithm. Points where the algorithm converges to a feasible optimum are members of $\mathcal{S}$ whereas points where the algorithm diverges are not. The expansion of $\mathcal{S}$ resulting from increasing n_b is colored in orange. Increments of n_b rapidly reduce the offset between Θ and $\mathcal{S}(n_b)$. This offset occurs near the boundary of Θ in all cases. This was seen in all numerical experiments performed. Note that the geometry of the staircase-infeasible domain falling within Θ (e.g., the region in red corresponding to $n_b = 100$) depends on the boundary of the set. For instance, the top left plot shows that increasing n_b only expanded the set of variables that are staircase-feasible over the nonlinear boundary. Whereas most of Θ is also in $\mathcal{S}(100)$, that is not the case for most standard families of random variables, e.g., a 4-parameter Beta with support Δ_z can only realize limited combinations of μ, m_2, m_3, and m_4 and the representable values of a moment depend upon the values taken by other moments. The bottom right plot shows that both staircase-feasible sets have holes and icicle-looking spikes. These features, which are caused by changes in the width of the bins that result from changing Ω, disappear when a fixed number of bins per unit of length, ρ, is used, i.e., $n_b = \lceil \rho(\overline{z} - \underline{z}) \rceil$. Means for seeking a staircase variable that closely approximates a staircase infeasible h are available in [20].

[3] The developments that follow are applicable whether the constraint in (24) is considered or not. However, for simplicity in the presentation we assume the latter case.

6. Staircase estimation

Suppose samples $z^{(1)}, \ldots, z^{(N)}$ are drawn independently from a fixed, but otherwise unknown DGM. This section focuses on the estimation of the hyper-parameter $h = [\theta_z, n_b]$ of a staircase variable S_z given data. Expert opinion should be used to prescribe n_b and the bound, Ω_z, to the support. Whereas Ω_z must[4] contain $\dot{\Delta} = [\min\{z^{(j)}\}, \max\{z^{(j)}\}]$, the moments μ, m_2, m_3, and m_4 can be chosen to be the sampling moments $\dot{\mu}$, $\dot{m}_2$, $\dot{m}_3$ and $\dot{m}_4$. The developments that follow account for the error incurred by using these estimates.

Lets first focus on the error incurred by using $\Omega_z \supset \dot{\Delta}$. Finite values of N often make $\dot{\Delta}$ a subset of the support set Δ. Scenario optimization [21] enables bounding the probability of the true distribution of z extending beyond $\dot{\Delta}$. In particular,

$$\mathbb{P}\left[z \notin \dot{\Delta}\right] \leq \hat{\epsilon}, \tag{14}$$

where

$$\hat{\epsilon} = 1 - e^{\log(\beta)/(N-1)}, \tag{15}$$

and β is the confidence parameter. See the proof in the Appendix. Eq. (14) is a formally verifiable, distribution-free and non-asymptotic result applicable to any stationary DGM. Scenario optimization theory states that Eq. (14) holds with probability greater than $1-\beta$, where β can be made very small such that it looses any practical significance. This probability is key for obtaining guaranteed results that are independent of the DGM. In determining a suitable staircase variable, Eq. (14) is equivalent to

$$\kappa q^{\top} \ell \leq \hat{\epsilon}, \tag{16}$$

where, q_i is the fraction of the interval $(z_i, z_{i+1}]$ falling outside $\dot{\Delta}$. This constraint ensures that the staircase variable conforms to (14).

We now focus on the error incurred by using the sample moments in θ. This error, prescribed by the corresponding sampling distributions, can be quantified using bootstrapping techniques or the central limit theory. This latter approach leads to the normal sampling distributions:

$$\mu \sim \mathcal{N}_{\mu}\left(\dot{\mu}, \sqrt{\frac{\dot{m}_2}{N}}\right),$$

$$m_2 \sim \mathcal{N}_{m_2}\left(\dot{m}_2, \sqrt{\frac{\dot{m}_4 - \dot{m}_2^2}{N}}\right),$$

$$m_3 \sim \mathcal{N}_{m_3}\left(\dot{m}_3, \sqrt{\frac{\dot{m}_6 - \dot{m}_3^2 - 6\dot{m}_4\dot{m}_2 + 9\dot{m}_2^3}{N}}\right),$$

$$m_4 \sim \mathcal{N}_{m_4}\left(\dot{m}_4, \sqrt{\frac{\dot{m}_8 - \dot{m}_4^2 - 8\dot{m}_5\dot{m}_3 + 16\dot{m}_2\dot{m}_3^2}{N}}\right),$$

conditional on $\theta \in \Theta$, where $\dot{m}_k$ is the sample moment of order k. These expressions correspond to an arbitrarily distributed variable z for a sufficiently large value of N [22]. For small values of N, bootstrapping techniques often yield a more accurate approximation. To account for sampling error in the estimation of a staircase variable, the moment matching constraints in (8) are replaced by the polynomial inequality constraints

$$\underline{\mu} \leq \mu(\ell) \leq \overline{\mu}, \tag{17}$$

$$\underline{m}_2 \leq m_2(\ell) \leq \overline{m}_2, \tag{18}$$

$$\underline{m}_3 \leq m_3(\ell) \leq \overline{m}_3, \tag{19}$$

$$\underline{m}_4 \leq m_4(\ell) \leq \overline{m}_4, \tag{20}$$

where $\mu(\ell) = r_2\ell$, $m_2(\ell) = r_3\ell - \mu(\ell)^2$, $m_3(\ell) = r_4\ell - \mu(\ell)^3 - 3\mu(\ell)m_2(\ell)$ and $m_4(\ell) = r_5\ell - 4\mu(\ell)m_3(\ell) - 6\mu(\ell)^2 m_2(\ell) - \mu(\ell)^4$ are the moments of the staircase, r_i is the ith row vector of A in (9), and the moment bounds correspond to the $1-\alpha$ confidence intervals of the sampling distribution. For instance, $\dot{\mu} - 1.96\sqrt{\dot{m}_2/N} \leq \mu(\ell) \leq \dot{\mu} + 1.96\sqrt{\dot{m}_2/N}$ for a 95% confidence interval. Note that the box of moments defined by (17)–(20) might not be fully contained in Θ.

Hence, the effects of sampling error are accounted for by solving for a staircase subject to the inequality constraints (14) and (17)–(20). Note however that the resulting variable will not account for the manner in which the sampling

[4] Sampling estimates will be denoted with a dot-superscript.

distribution allocates probability within the chosen box of moments. This consideration can be taken into account by using the cost

$$J(\ell) = -E(\ell) - \log\{L(\ell)\}, \tag{21}$$

where $L = \mathcal{N}_{\mu(\ell)}\mathcal{N}_{m_2(\ell)}\mathcal{N}_{m_3(\ell)}\mathcal{N}_{m_4(\ell)}$ is the likelihood function of the sampling distribution for independent moments.

In conclusion, staircase variables provide (i) the ability to represent a wide range of density shapes, (ii) the ability to represent most of the feasible space Θ, (iii) the framework to account for the sampling error present during their estimation from data, and (iv) the low-computational cost required to efficiently perform many uncertainty quantification tasks. This cost refers to both the calculation of the density function and the simulation of the resulting staircase variable.

7. Examples

Staircase variables are used below to generate empirical predictor models. To this end we assume that N independent and identically distributed input-output pairs are obtained from a Data Generating Mechanism (DGM), and denote by $\mathbb{D} = \{x^{(i)}, y^{(i)}\}$, $i = 1, \ldots N$, the corresponding data sequence. For instance, in the modal analysis of a flexible structure, the input x is a frequency of excitation whereas the output y is the corresponding amplitude of the response at a particular sensor location. Random variations in material properties, boundary conditions and measurement noise make the output y aleatory.

The models of interest are called *Staircase Predictor Models* (SPM) because they characterize the output for any given input as a staircase variable. The SPMs presented below are data-based. As such, they don't require making any assumption on the manner by which the data in $\mathbb{D}$ is distributed. SPMs require prescribing the input-dependent θ functions

$$\theta_{y(x)} = \left[\underline{y}(x), \overline{y}(x), \mu_{y(x)}, m_{2,y(x)}, m_{3,y(x)}, m_{4,y(x)}\right], \tag{22}$$

over the input range X. Note that $\theta_{y(x)}$ is a functional extension of θ_z in Eq. (2). The technique for prescribing the target functions $\theta_{y(x)}$ according to $\mathbb{D}$ proposed in [23] is used in the examples below.

Example 1: Model verification. In this example we generate a SPM for a DGM about which we have full knowledge. As in the method of manufactured solutions, model verification is performed by comparing the resulting model predictions with the true underlying phenomenon being modeled.

To this end we consider the DGM shown at the top of Fig. 4. The probability of falling between any pair of adjacent lines is 0.01. Note that the distribution, including its support, location and number of modes, varies strongly with the input. A data sequence[5] with $N = 1000$ observations was drawn from the DGM. Fig. 5 shows the functions corresponding to the DGM (solid lines) along with the target functions $\theta_{y(x)}$ (dashed lines) extracted from the data. Note that the target functions approximate the DGM functions reasonably well in spite of only using 1000 observations.

The target functions $\theta_{y(x)}$ were then used to build the *moment-matching SPM* $S_{y(x)}(\theta_{y(x)}, 500, E)$. The entropy of this predictor is $\mathbb{E}_x[E] = -0.3642$. Fig. 4 shows this SPM at the second subplot. This figure was generated by calculating staircase variables over a uniform grid of input values in X, sampling them, and grouping the points belonging to the same percentile line. The moment functions attained by the SPM are indistinguishable from the targets. The comparison between the DGM and the SPM indicates an excellent agreement, as measured by the squared difference between pairs of corresponding percentiles. Note that the SPM describes well the bimodal structure of the DGM by replicating the regions where probability is highly concentrated, i.e., the regions in the upper and lower limit of the support where many percentile lines cluster. Furthermore, the skewness of the probability mass in the interior of the support set follows the same patterns present in the DGM. All this information is implicitly embedded in $\theta_{y(x)}$. The comparison of this SPM against alternative modeling techniques is available in Reference [23].

Next we evaluate the effects of sampling error in the prescription of $\theta_{y(x)}$. To quantify the sparsity of the dataset we define the equivalent number of observations, n_e, as

$$n_e(x) = \sum_{j=1}^{N} w(x^{(j)}, x),$$

where $0 \le w \le 1$ is the weighting function used in [23]. The sampling error in the moment estimates will be quantified by using the developments of Section 6 for $N = n_e(x)$. For this data sequence, the value of n_e ranges between 47.15 and 112. This indicates that the dataset is sparse. Fig. 6 shows the 95% confidence intervals of the sampling error for the four moment functions as shaded regions. Note that the interval functions oscillate considerably, reaching their largest spread near $x = 0$.

Next we calculate a maximal-entropy SPM whose moment functions are free to take on any value within these interval functions. Hence, each staircase variable comprising the SPM will maximize the entropy while satisfying the spread constraint (14), along with the box-of-moments constraints (17–20). We will refer to this predictor as a *moment-bounded SPM*. The moments attained by this SPM are shown as solid lines in Fig. 6. The first three moments vary in the interior of their intervals whereas the fourth moment stays on the lower limit. The relaxation of the moment-matching constraints to moment-bounded constraints increases the expected entropy from $\mathbb{E}_x[E] = -0.3642$ to $\mathbb{E}_x[E] = 0.7484$. The resulting SPM is shown in the third subplot of Fig. 4. Note that the most prominent features of the random process, such as the peaks at the

[5] The data sequence and numerical setup used here were also used in an example of reference [24]. Additional details can be found therein.

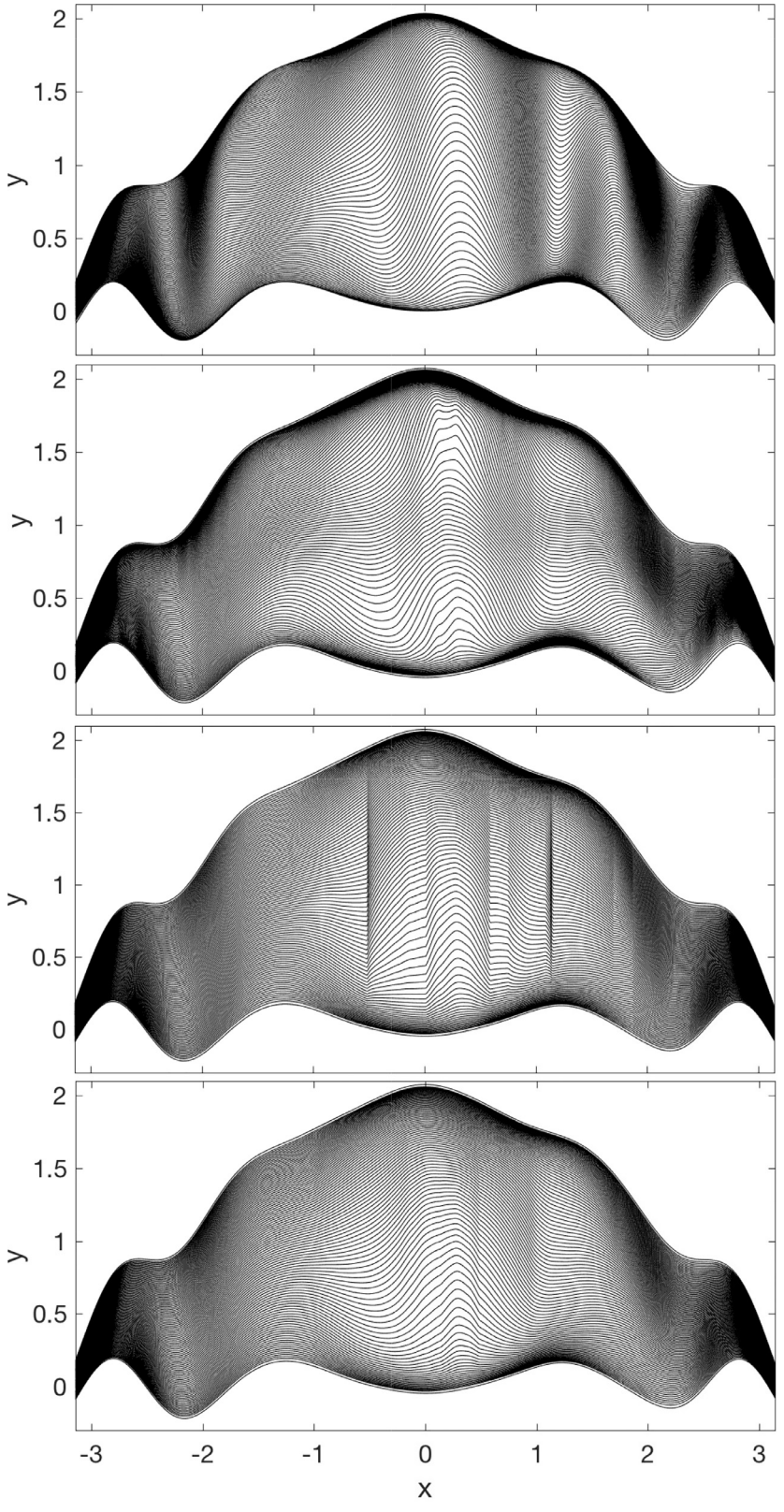

Fig. 4. DGM (top), maximal-entropy moment-matching SPM (second), maximal-entropy moment-bounding SPM (third), and minimal-(21) moment-bounding SPM (bottom).

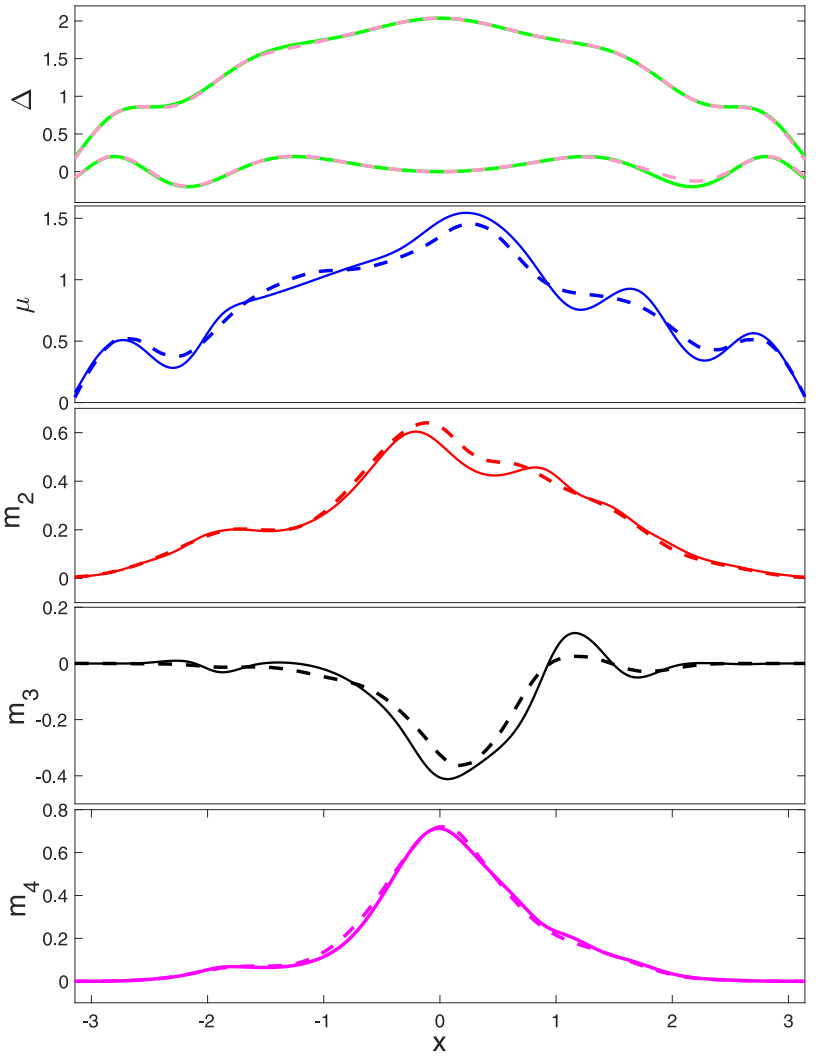

Fig. 5. Δ_z and moments for the DGM (solid lines) and the target (dashed lines).

boundaries of the support set, and the patterns of the lines in its interior, have faded. Furthermore, note that the derivative discontinuities in the moments functions of Fig. 6 yield derivative discontinuities in the percentile lines at the corresponding input values. Such discontinuities can be eliminated by using a kernel smoother [25] or by using another cost function.

The latter approach is considered next. In particular, we will calculate a moment-bounded SPM having the cost function in Eq. (21). By using this cost the resulting staircase variables account for the manner in which the sampling distribution allocates probability within the box of moments. The moment functions corresponding to this SPM are shown as dotted lines in Fig. 6. In contrast to the other moment-bounded SPM, the new moments have continuous derivatives throughout X. The resulting SPM, shown at the bottom of Fig. 4, exhibits smooth percentile lines. This is achieved at the expense of a minor entropy reduction from $\mathbb{E}_x[E] = 0.7484$ to $\mathbb{E}_x[E] = 0.7300$. As n_e increases, the width of the confidence intervals reduces, and the moment-bounded SPMs converge to the moment-matching SPM. When the value of n_e is sufficiently large, the dataset will no longer be sparse, and the analyst might prefer using target-matching or maximum log-likelihood staircases.

*Example 2: Flutter analysis.*Next we study the stability of an airfoil with pitch and plunge degrees of freedom subject to unsteady aeroelastic effects [26]. Flutter, a non-linear response caused by a Hopf bifurcation in the system dynamics, might occur when the free stream speed increases. Therein, energy in the flow is transferred to the airfoil causing a self-sustained oscillation that can potentially lead to loss of control and structural damage. The airflow speed at which this occur, called the flutter speed, is a complex function of the inertial, geometrical, material, and aerodynamic properties, and is known to be highly sensitive to uncertainties [27]. The physics-based model acting as the DGM is presented in the appendix. This model, an abstracted version of a flexible three-dimensional wing model with bending and torsional structural dynamics [26], is commonly used for control system analysis of actual fixed-wing aircraft.

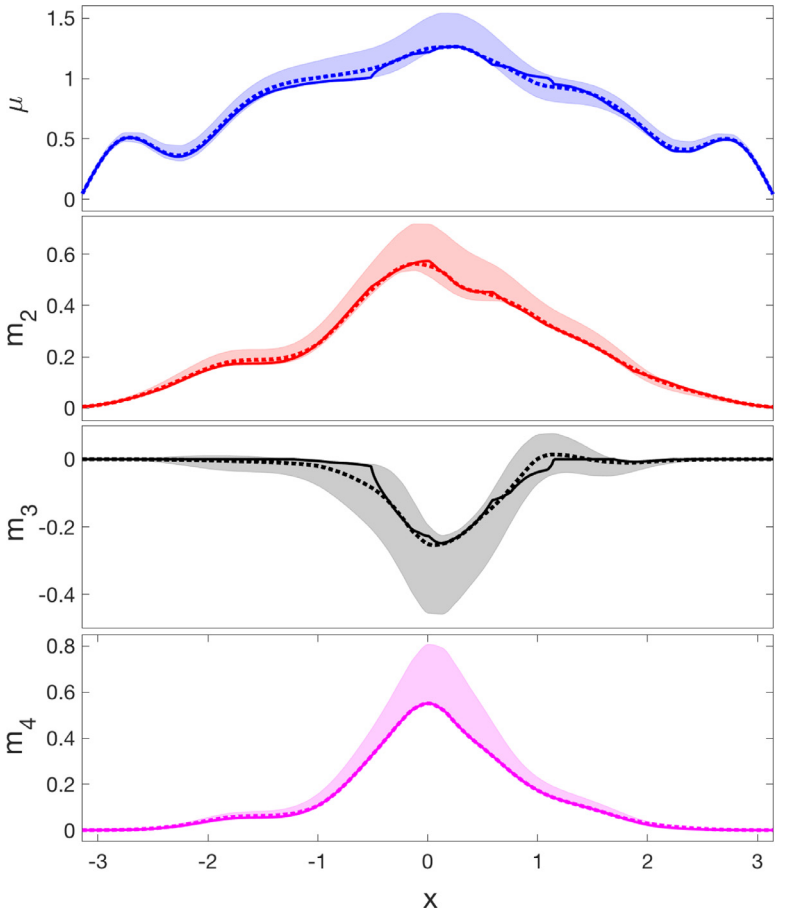

Fig. 6. Moments of a moment-bounded SPM based on (11) (solid line), of a moment-bounded SPM based on (21) (dotted line), and the corresponding bounds (shaded regions).

The dynamic response to a particular airflow speed x (i.e., the input in the context of this paper, denoted as V in the appendix) is evaluated by calculating the stability parameter, y (i.e., the output in the context of this paper). Non-negative values of y denote an unstable response, whereas negatives values correspond to a stable response. Variability in the system parameters makes y a random process. We not only want to determine the probability of instability as a function of the airflow speed, $P[y(x) > 0]$, but also the manner by which the onset of instability occurs. Instabilities for which y is positive but not too large (so the time for the amplitude of the response to double is sufficiently large) can be mitigated by using a flight control system or by reducing the speed, whereas instabilities for which y is positive and large might not be recoverable.

As in the previous example, an SPM will be used to characterize the system response given data. To that end, the data sequence $\mathbb{D} = \{x^{(i)}, y^{(i)}\}$ for $i = 1, \ldots 14000$, was obtained by simulating the model in the[6] Appendix A.3. Fig. 7 shows the data. Note that the range and distribution of the data vary strongly with the airflow speed x. The range varies from a practically zero spread at low values of x to sizable spreads at high values of x. Furthermore, note that the transition from stability to instability occurs in the range $x \in [1.93, 2.71]$, such that $P[y(1.93) > 0] = 0$ and $P[y(2.71) > 0] = 1$.

The data sequence was then used to estimate the target functions in Eq. (22). Fig. 7 shows the bounds to the range of the data whereas Fig. 8 shows the target moment functions and their corresponding sampling error intervals. Note that the sampling error is negligible when $x < 3$, and it becomes larger as x increases. The third-order moment function shows that the skewness of the distribution varies considerably taking on values that are zero, positive and negative. Since the densities of the uncertainties prescribing the random response are symmetric, the skewness in the response is solely caused by the underlying dynamics. The targets were then used to calculate a moment-matching staircase predictor. Fig. 9 shows the two-percentile curves of the resulting SPM.

[6] Making the DGM a realistic computational model is an intended application setting for the proposed methods. The main reason for this is that "real data" is often very expensive or impossible to obtain (e.g. the prediction of performance of a structure in a zero-gravity environment). Studying the effects of uncertainty in the predictions resulting from computational models is a subject of paramount importance when assessing their performance, safety and reliability.

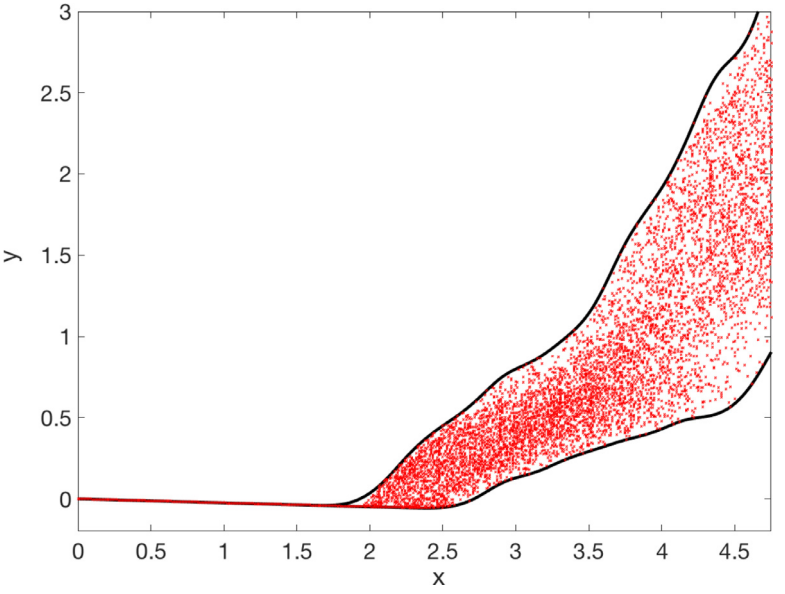

Fig. 7. Data ($\times$), and bound to the support set (lines).

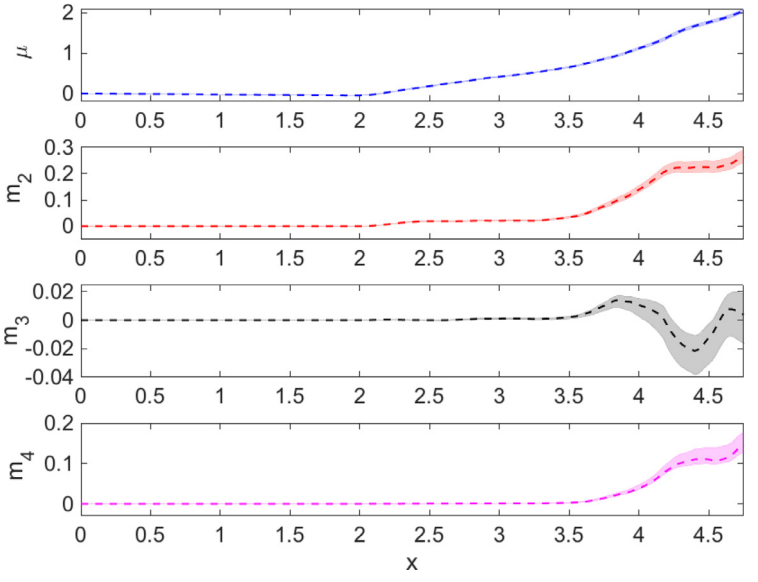

Fig. 8. Target functions (dashed lines) and their sampling error ranges (shaded areas).

Note that this model accurately captures the main features of the data including the input-dependent range and the variations in the skewness. It is worth noticing that the large variability observed in the system response is caused by small parametric uncertainties of less than 5 % (see Appendix).

For comparison sake, a Heteroscedastic Gaussian Process (GP) model and a Gaussian Mixture (GM) model based on the same data sequence $\mathbb{D}$ were generated. The corresponding 2-percentiles are shown in Fig. 10. Note that the GP model fails to describe the practically deterministic response near $x = 0$, and that the prediction in the domain where the transition to instability occurs, i.e., $2 < x < 3$, differs substantially from the data. More importantly, the inability of the model to yield non-Gaussian responses often make the resulting predictions unacceptable. In the context of this example this implies obtaining a spurious prediction of the probability of instability throughout the input range. The GM[7] model on the right of Fig. 10 yields a better representation of the data. The number of model's components as well as bounds to the covariance

[7] This model was obtained by first characterizing the joint density of the input and output, and then calculating the conditional density of the output for a grid of inputs. This model was calculated by using the expectation-maximization algorithm with 200 components.

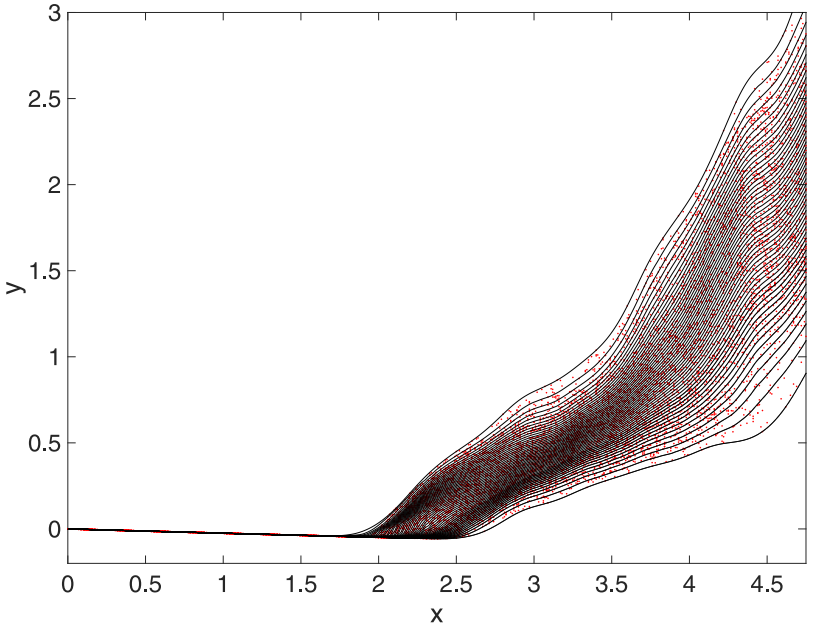

Fig. 9. 2-percentiles of the SPM $S_{y(x)}(\theta_{y(x)}, 250, E)$.

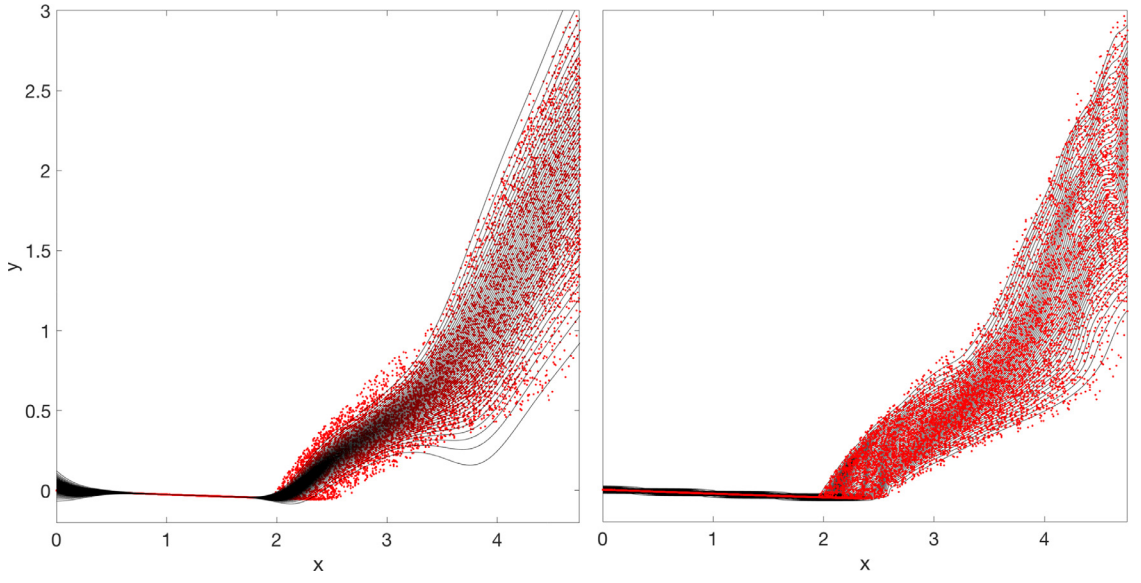

Fig. 10. Heteroscedastic GP model (left) and a GM model (right).

matrices were tuned to prevent obtaining an non-smooth predictor with ill-conditioned covariance matrices. Even though the GM model describes the data reasonably well, the predicted distribution in $x < 2$ has an overly large variance, and several spurious patterns occur at the interior of the distribution, e.g., see a spurious oscillation of most percentiles near $x = 4.6$. GM models face challenges when describing densities varying almost linearly with the input, densities with discontinuities, e.g., a uniform random variable and the DGM in Example 1, and densities with small variances. The characterization of these features requires a large number of components with small covariance matrices, making the predictor non-smooth and the corresponding optimization ill-conditioned. The comparison of the models in Figs. 9 and 10 relative to the data indicates that the SPM is the best modeling choice. The SPM not only describes the distribution of the data more accurately than the two alternatives but also avoids the significant amount of tuning required by them.

Next, the SPM will be used to evaluate the probability of instability as a function of airspeed, as well as the severity of such an instability, i.e., the degree of difficulty required to fly the aircraft back to a stable regime. To formalize this notion,

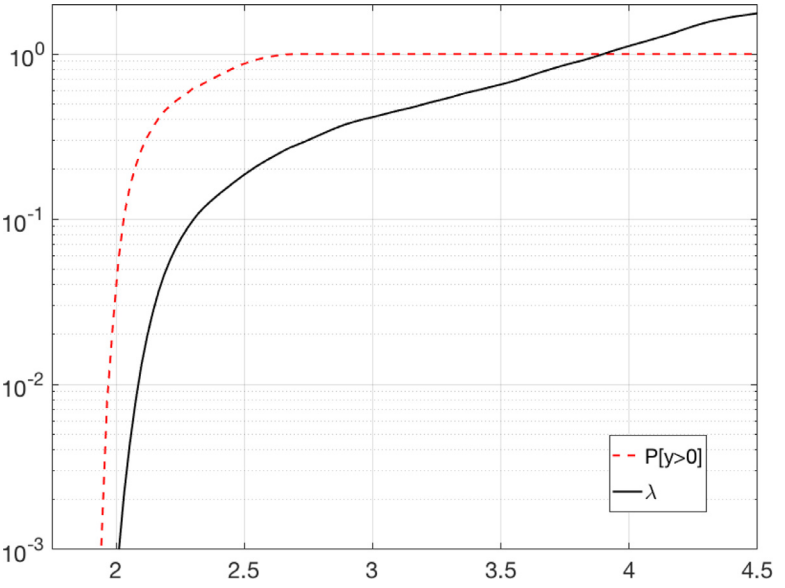

Fig. 11. Instability metrics as a function of airspeed.

consider the metric

$$\lambda(x) := \int_0^{\bar{y}(x)} y(x) f_{y(x)}(y) dy. \tag{23}$$

When $\lambda = 0$ the probability of instability is zero. When $\lambda > 0$ the fraction of unstable responses contributes to λ proportionally to the value of y. Note that λ might not be proportional to the probability of instability; e.g., a response with a probability of instability of 0.1 and a long upper tail can achieve a greater value of λ than a response with a probability of instability of one. Fig. 11 shows the probability of instability $P[y > 0]$ and λ as functions of the airspeed. The admissible range of airflow speeds for zero probability of instability is $x \in [0, 1.93]$. If the set of responses for which $\lambda < 0.1$ is considered safe/recoverable, airspeeds below $x = 2.3$ become admissible. Note that probabilities of instability as large as 0.62 will occur within this range. Therefore, the range of airflow speeds is expanded 20% by increasing the admissible value of λ from zero, which corresponds to a zero probability of flutter instability, to 0.1.

The staircase predictor above was instrumental in accurately prescribing a safe flying range while accounting for the severity of the instability.

8. Conclusions

This paper proposes moment-matching staircase variables for uncertainty modeling. These variables have a bounded support set, and prescribed values for the first four moments. We propose a convex optimization framework for calculating their distributions according to one of several optimality criteria. These criteria include maximal entropy, minimal squared amplitude, optimal target matching, and maximal log-likelihood. Convexity enables substantial computational advantages including fast computational times, and the ability to solve for practically smooth probability densities using interior point methods. This feature makes staircases well suited for many practical applications requiring their repeated calculation and simulation, e.g., the Bayesian calibration of a computational model whose parameters are staircases. In addition, staircases enable modeling phenomena having a possibly skewed and/or multi-modal response varying over a bounded range.

Modeling a random process whose support set and distribution vary strongly with the input variable, as in both of the paper's examples, is often a challenge. Gaussian process models, the most widely used non-parametric metamodeling technique, fail to accurately describe many phenomena due to their intrinsic structure. Gaussian mixture models are better suited for modeling complex responses. However, they face challenges when the underlying density varies (almost) linearly with the input, when it has discontinuities, or when it has small variances. Staircase predictor models do not suffer from such deficiencies nor do they require the significant amount of tuning required by such models. This tuning it is required to not only set the model structure and its parameters from many available options, but also to avoid the numerical search for the optimum of a non-convex program from being trapped in local minimum.

Furthermore, we propose a strategy for staircase density estimation given data. This approach seeks a staircase whose moments are free to take on any value within the prediction intervals of the moment estimates. The efficiency and flexibility of staircase variables is demonstrated by generating empirical predictor models for two application examples. The resulting staircase predictors not only accurately describe the complex responses of the underlying phenomena throughout the input

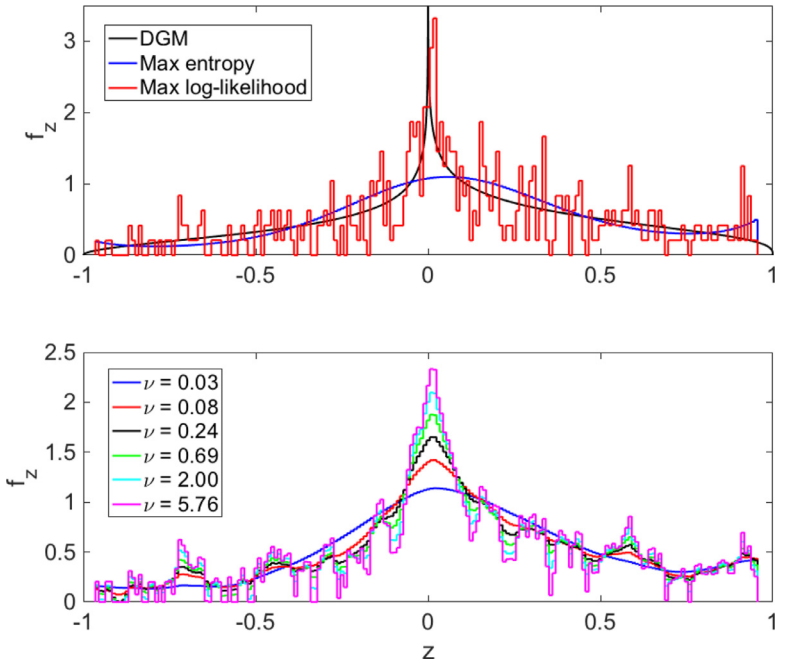

Fig. 12. Unrestricted (top) and restricted maximal likelihood staircases (bottom).

range, but also are instrumental in performing a joint risk and uncertainty analysis for safe flight. Future work will focus on further constraining staircase variables so unimodality, symmetry, smoothness and connectedness considerations can be enforced within the convex programming framework.

Appendix A

A1. Jump attenuation

The density of maximal likelihood staircase variables might exhibit large jumps between neighboring bins. Even though the flexibility of a staircase variable enables achieving likelihoods that far exceed those attained by variables from other families, so they would better describe the data, one might want to attenuate such jumps. This can be attained by further constraining (7) with the convex constraint

$$\ell^\top C \ell < \nu, \tag{24}$$

where $\nu > 0$ is inversely proportional to the desired level of smoothness, and $C \in \mathbb{R}^{n_b \times n_b}$ satisfies

$$C = \begin{bmatrix} 1 & -1 & & & & \\ -1 & 2 & -1 & & 0 & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & 0 & & -1 & 2 & -1 \\ & & & & -1 & 1 \end{bmatrix} \succeq 0. \tag{25}$$

The constraint in (24) bounds the sum of the square difference of the likelihood at adjacent bins. The top of Fig. 12 displays the maximal likelihood variable corresponding to $N = 500$ observations drawn from the DGM in Fig. 2. Densities for the DGM and the maximal entropy are superimposed. The jumps are the result of large differences in the number of observations falling in neighboring bins. The bottom of Fig. 12 shows the staircase variables that result from solving (8) with cost (13) and constraint (24) for several values of ν. Note that the parameter ν controls the degree of smoothness of the resulting density. The maximal likelihood variable at the top is obtained when ν is large.

A2. Probability bound

The proof of Eq. (14) is as follows. The tightest interval containing the samples, $\dot{\Delta}$, is the solution of the convex program $\min_{a,b}\{b - a : a \le z^{(j)}, b \ge z^{(j)}, j = 1, \dots N\}$. As such, scenario optimization theory can be readily applied to bound the

probability of future samples falling outside $\dot{\Delta}$. Theorem 2.1 in [28] for $k = 0$, and $d = 2$, leads to

$$\hat{\epsilon} = \underset{\epsilon}{\text{argmax}}\{\epsilon : (1-\epsilon)^N + N\epsilon(1-\epsilon)^{N-1} \leq \beta\},$$

whose optimum is Eq. (15).

A3. Dynamical model for flutter analysis

The flutter equations of an elastically-mounted airfoil [26] are

$$\begin{bmatrix} m & mbx_\theta \\ mbx_\theta & I \end{bmatrix}\begin{bmatrix} \ddot{h} \\ \ddot{\theta} \end{bmatrix} + \begin{bmatrix} k_h & 0 \\ 0 & k_\theta \end{bmatrix}\begin{bmatrix} h \\ \theta \end{bmatrix} = \begin{bmatrix} -L \\ b(a+1/2)L + M \end{bmatrix}, \tag{26}$$

where h is the plunge motion, θ is the pitch motion, m is the mass of the airfoil, I is the moment of inertia, b is the airfoil's semi-chord, k_h is the plunge stiffness, k_θ is the pitch stiffness, L is the aerodynamic lift, and M is the aerodynamic pitching moment. The dimensionless parameters a and e define the equivalent location of the plunge spring connection and of the center of mass, respectively, $x_\theta = e - a$.

Unsteady aerodynamic loads are commonly modeled as

$$L = C_{L\alpha}\rho U b C(k)\left(\dot{h} + U\theta + b(1/2 - a)\dot{\theta}\right) + \pi\rho b^2\left(\ddot{h} + U\dot{\theta} - ab\ddot{\theta}\right), \tag{27}$$

$$M = -\pi\rho b^3\left(\ddot{h}/2 + U\dot{\theta} + b(1/8 - a/2)\ddot{\theta}\right), \tag{28}$$

where $C_{L\alpha}$ is the lift slope with respect to angle of attack, ρ is the flow density, U is the flow speed, and $C(k)$ is the Theodorsen function [29] written in terms of the reduced frequency k. Unsteady aerodynamic effects are often modeled using the Theodorsen's function, which is a complex-valued function of reduced frequency meant to capture both the amplitude attenuation and the phase lag of unsteady air loads. This function makes the dynamics in Eqs. (26)–(28) nonlinear. In order to solve this equation, a solution of the following form is assumed:

$$\begin{Bmatrix} h \\ \theta \end{Bmatrix} = \boldsymbol{x} = \bar{\boldsymbol{x}}\, e^{\,pU/b}, \tag{29}$$

where p is defined as $p = g + k\,i$, k is the aforementioned reduced frequency, and g is a damping parameter, and i is the imaginary unit. Combining (26)–(29) results in:

$$\left[\left(U^2/b^2\right)\boldsymbol{M}\,p^2 + \boldsymbol{K} + q\,\boldsymbol{A}(p)\right]\bar{\boldsymbol{x}} = \boldsymbol{0} \tag{30}$$

where $q = \rho U^2/2$ is the dynamic pressure, $\boldsymbol{M}$ and $\boldsymbol{K}$ are the structural matrices from (26), and the remaining terms are collected in $\boldsymbol{A}$, a complex-valued aerodynamic matrix which is a highly nonlinear function of p both due to the time-derivatives of h and θ, which lead to p and p^2 terms, but also due to the dependence of Theodorsen's function on the imaginary part of p.

The resulting nonlinear eigenvalue equation may be solved by selecting an initial guess for each eigen-pair $(p, \bar{\boldsymbol{x}})$ and iterating to convergence, a process commonly known as the *p-k method* [30]. Alternatively, Eq. (30) may be converted into a larger linear eigenvalue equation by using the rational function approximation:

$$\boldsymbol{A}(p) \approx \boldsymbol{A}_0 + \boldsymbol{A}_1 p + \boldsymbol{A}_2 p^2 + \sum_{j=3}^{n} \frac{\boldsymbol{A}_j p}{p + \gamma_{j-2}} \tag{31}$$

where γ_j are aerodynamic lag roots placed at strategic reduced frequency k values. Since $\boldsymbol{A}$ is a known function of $k\,i$, the $\boldsymbol{A}_j$ terms may be computed via least-squares along the imaginary axis.

$$\boldsymbol{A}_r(k\,i) + i\boldsymbol{A}_i(k\,i) \approx \boldsymbol{A}_0 + \boldsymbol{A}_1\,\boldsymbol{k}\,i - \boldsymbol{A}_2 k^2 + \sum_{j=3}^{n} \frac{\boldsymbol{A}_j\,k\,i}{k\,i + \gamma_{j-2}} \tag{32}$$

Aerodynamic states $\boldsymbol{x}_a$ are designated as:

$$\boldsymbol{x}_{a_j} = \boldsymbol{x}\frac{p}{p + \gamma_{j-2}} \tag{33}$$

The combination of Eqs. (30), (31), and (33) leads to the Linear Parameter Varying (LPV) model in U:

$$(p\boldsymbol{S_1} + \boldsymbol{S_2})\begin{Bmatrix} \boldsymbol{x} \\ p\boldsymbol{x} \\ \boldsymbol{x}_{a_1} \\ \vdots \\ \boldsymbol{x}_{a_n} \end{Bmatrix} = \boldsymbol{0}, \tag{34}$$

where

$$\boldsymbol{S_1} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \frac{U^2}{b^2}\boldsymbol{M} + q\boldsymbol{A}_2 & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{I} \end{bmatrix},$$

$$\boldsymbol{S_2} = \begin{bmatrix} \boldsymbol{0} & -\boldsymbol{I} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{K} + q\boldsymbol{A}_0 & q\boldsymbol{A}_1 & q\boldsymbol{A}_3 & \cdots & q\boldsymbol{A}_{n+2} \\ \boldsymbol{0} & -\boldsymbol{I} & \gamma_1\boldsymbol{I} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & -\boldsymbol{I} & \boldsymbol{0} & \cdots & \gamma_n\boldsymbol{I} \end{bmatrix}.$$

This approach is called the Roger's approximation [31] or the p-method.

This model can be written in terms of six nondimensional quantities, a, e, $\mu = m/(\pi\rho b^2)$, $r = \sqrt{I/(mb^2)}$, $\sigma = \omega_h/\omega_\theta$ and $C_{L\alpha}$ where the natural frequencies are defined as $\omega_h = \sqrt{k_h/m}$ and $\omega_\theta = \sqrt{k_\theta/I}$. Nominal values for this parameters are $a = -0.2$, $e = -0.1$, $\mu = 20$, $r = 0.42$, $\sigma = 0.4$ and $C_{L\alpha} = 2\pi$. Furthermore, a non-dimensional flow speed is defined as $V = U/b/\omega_\theta$. With the LPV model in (34) at hand, the system stability is prescribed by the manner in which the corresponding eigenvalues vary with V. In the example above we assume that $x = V$ and y is the maximal real part of the plunge, pitch, and lag eigenvalues. It is further assumed that the model response is aleatory because the parameters a and e vary uniformly within 5% from their nominal value.

References

[1] T. Simpson, J. Peplinski, P. Koch, J. Allen, Metamodels for computer-based engineering design: survey and recommendations, Eng. Comput. 17 (1) (2001) 129–150.
[2] M. Allen, K. Maute, Reliability-based design optimization of aeroelastic structures, Struct. Multidiscip. Optim. 27 (2004) 228–242.
[3] M.S. Eldred, H. Agarwal, V. Perez, S. Wojtkiewicz, J. Renaud, Investigation of reliability method formulations in DAKOTA/UQ, Struct. Infrastruct. Eng. Maint. Manag. Life Cycle Des. Perform. 3 (3) (2007) 199–213.
[4] L.P. Swiler, B. Adams, M. Eldred, Model calibration under uncertainty: Matching distribution information, in: Proceedings of the AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, Victoria, British Columbia, Canada, 2008.
[5] C.C. McAndrew, Compact modeling: principles, techniques, and applications, Statistical Modeling using Backward Propagation of Variance, Springer, New York, 2010.
[6] D. Bertsimas, I. Popescu, On the relation between option and stock prices: an optimization approach, Oper. Res. 50 (2) (2002) 358–374.
[7] B. Grundy, Option prices and the underlying asset's return distribution, J. Finance 46 (3) (1991) 343–556.
[8] L.E. Ghaoui, M. Oks, F. Oustry, Worst-case value-at-risk and robust portfolio optimization: a conic programming approach, Oper. Res. 51 (4) (2003) 358–374.
[9] J. Smith, Generalized chebyshev inequalities: theory and applications in decision analysis, Oper. Res. 43 (5) (1995) 358–374.
[10] R. Sharma, S. Devi, G. Kapoor, N.S. Barnett, A brief note on some bounds connecting lower order moments for random variables defined on a finite interval, Int. J. Theor. Appl. Sci. 1 (2) (2009) 83–85.
[11] R. Sharma, R. Kumar, R. Saini, Kapoor, Complementary upper bounds for fourth central moment with extensions and applications, arXiv:1503.03786 (2015).
[12] P. Kumar, Moment inequalities of a random variable defined over a finite interval, J. Inequ. Pure Appl. Math. 3 (3, Article 41) (2002) 1–11.
[13] J. Smith, Moment methods for decision analysis, Ph.D. Thesis, Stanford University(1990).
[14] D. Bertsimas, I. Popescu, Optimal inequalities in probability theory: a convex optimization approach, SIAM J. Optim. 15 (3) (2005) 780–804.
[15] I. Popescu, A semidefinite programming approach to optimal moment bounds for convex classes of distributions, Math. Oper. Res. 30 (1) (2005) 1–23.
[16] R. Tian, S. Cox, L.F. Zuluaga, Moment problem and its applications to risk assessment, North Am. Actuarial J. 21 (1) (2017) 242–266.
[17] A.S. Nemirovski, M.J. Todd, An approximate true damping solution of the flutter equation by determinant iteration, Acta Numerica 1 (2008) 191–234.
[18] B.W. Silverman, Density Estimation for Statistics and Data Analysis, Chapman and Hall, 11 New Fetter Lane, London, England, 1986.
[19] C. Fraley, A.E. Raftery, Model-based clustering, discriminant analysis, and density estimation, J. Am. Stat. Assoc. 97 (458) (2002) 611–631, doi:10.1198/016214502760047131.
[20] L.G. Crespo, D.P. Giesy, S.P. Kenny, On the calculation and shaping of staircase random variables, in: ESREL 2017, Portoroz, Slovenia, 2017.
[21] M.C. Campi, S. Garatti, The exact feasibility of randomized solutions of uncertain convex programs, SIAM J. Optim. 19 (3) (2008) 1211–1230.
[22] M. Kendall, A. Stuart, The Advanced Theory of Statistics, third ed., Charles, Griffin and Co., London, 1969.
[23] L.G. Crespo, S.P. Kenny, D.P. Giesy, Staircase predictor models for reliability and risk analysis, Structural Safety 75 (2018) 35–44.
[24] L.G. Crespo, D.P. Giesy, S.P. Kenny, Random predictor models with a nonparametric staircase structure, in: ESREL 2017, Portoroz, Slovenia, 2017.
[25] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning, second ed., Springer, New York, 2001.
[26] D. Hodges, A. Pierce, Introduction to Structural Dynamics and Aeroelasticity, Cambridge University Press, New York, 2002.
[27] P. Beran, B. Stanford, C. Schrock, Uncertainty quantification in aeroelasticity, Ann. Rev. Fluid Mech. 49 (1) (2017) 361–386.

[28] M.C. Campi, S. Garatti, A sampling-and-discarding approach to chance-constrained optimization: feasibility and optimality, J. Optim. Theory Appl. 148 (2) (2011) 257–280.
[29] T. Theodorsen, General Theory of Aerodynamic Instability and the Mechanism of Flutter, NACA, New York, 1949. report 496
[30] H. Hassig, An approximate true damping solution of the flutter equation by determinant iteration, J. Aircraft 8 (11) (1971) 885–889.
[31] K. Roger, Airplane Math Modeling Methods for Active Control Design, Structural aspects of active control, AGARD Defense Technical Information Center DTIC AD A 045242 CP-228, Neuilly-Sur-Seine, France(1977) 4.1–4.11.